Oceanic vertical exchange and particle export: Comparisons between models and new observational estimates

Anand Gnanadesikan[1]
John P. Dunne[1]
Robert M. Key[2]
Katsumi Matsumoto[3]
Jorge L. Sarmiento[2]
Richard D. Slater[2]
P.S. Swathi[4]

submitted to Global Biogeochemical Cycles

1: NOAA/Geophysical Fluid Dynamics Lab, PO Box 308 Princeton, NJ 08542
2: AOS Program, Princeton University, PO Box CN710, Princeton, NJ 08540
3: Present address: AIST, Geological Survey of Japan, Tsukuba, Japan
4: Center for Mathematical Modelling and Computer Simulation, Bangalore, India

Abstract

This paper examines the extent to which oceanic vertical exchange in general circulation models can be constrained by biogeochemical measurements. In an earlier version of the Princeton Ocean Biogeochemical Model, the deep ocean was not well ventilated, and concentrations of radiocarbon and oxygen were too low, particularly in the deep North Pacific. Three solutions to this problem have been found: 1. correcting surface boundary conditions next to the Antarctic Continent, to ensure that Weddell Sea and Ross Sea deep waters actually outcrop during the wintertime 2. correcting surface boundary conditions throughout the Southern Ocean to better simulate the annual cycle of temperature and salinity, which increases vertical exchange in the open waters of the Southern Ocean. 3. increasing the vertical diffusion coefficient in the upper ocean above that which direct measurements of mixing seem to allow. Three models are presented that use different combinations of these solutions. The first corrects the boundary conditions near the Antarctic and moderately increases the vertical diffusion coefficient. The second applies corrections to the surface boundary conditions both around the Antarctic and in the open waters of the Southern Ocean. The third greatly increases the vertical diffusion coefficient. All three models produce more realistic oxygen and radiocarbon distributions. However, when the particle export predicted by the models is compared with that estimated using satellite color and an algorithm described in a companion paper, significant differences are found between the models. Implications of these differences are considered.

1. Introduction

Ever since the discovery in the late 18th century that deep waters at the equator are very cold, oceanographers have sought to understand the pathways by which deep abyssal and surface waters exchange with each other. There are largely two ways in which this exchange occurs. The first, concentrated in the tropics, is largely determined by the diapycnal mixing coefficient (Bryan, 1987). Increasing the diapycnal mixing coefficient results in more downward diffusion of light surface water and upwelling of dense deep water. The second, concentrated in the Southern Ocean, is largely determined by a competition between two processes. Southern Ocean winds drive upwelling, which can be either from deep dense waters or lighter waters from low latitudes. Mesoscale eddies tend to supply this upwelling from the low-latitude water. Gnanadesikan (1999a) proposed a simple theory linking the depth of the pycnocline to these three processes. This theory predicts that it is possible to alter the lateral and vertical mixing coefficients so as to preserve the shape of the thermocline and magnitude of Northern Hemisphere overturning, while altering which exchange pathway dominates. Gnanadesikan et al. (2002) showed that this was in fact the case. Klinger et al. (2003) verify that the theory works in both level-coordinate and isopycnal models.

The fact that temperature and salinity may be simulated reasonably well using different combinations of mixing coefficients suggests that other observations might be useful to better constrain the circulation. A possibility that we have been pursuing for many years is that measurements of biological cycling could provide such constraints. Biological cycling depends strongly on oceanic vertical exchange. Deep ocean waters are rich in inorganic nutrients such as nitrate and phosphate. Vertical exchange supplies

these waters to the surface ocean, where the nutrients are taken up by plankton, incorporated into organic matter, and exported to the deep ocean as dissolved or particulate organic matter. If we were able to accurately estimate either the rate of inorganic nutrient uptake, or the rate of particulate and dissolved organic matter export, we would be able to constrain the physical rate of vertical exchange, and even to draw conclusions about the physics driving ocean overturning. Based on the fact that relatively little siliceous sediment is found throughout the subtropical Pacific, Worthington (1977) argued that relatively little Antarctic Bottom Water is upwelled through the tropical thermocline. While, the lack of sediment could have been explained by significant dissolution within the water column, Gnanadesikan and Toggweiler (1999) showed that observed export from the mixed layer was also inconsistent with the strong upwelling of deep water which would be required to balance high levels of vertical mixing. Within the Southern Ocean, Gnanadesikan (1999b) argued that observational estimates of silica export from the mixed layer were much lower than those predicted using models in which the impact of eddy-induced transports (Gent and McWilliams, 1990) were neglected, but were consistent when these effects were included.

A weakness of the results of Gnanadesikan and Toggweiler (1999) and Gnanadesikan (1999b) is that they rely on direct estimates of particle export. This is a problem, since there are very few such estimates and these are sparse in space and time. Satellite estimates of ocean color, by contrast, have good spatiotemporal coverage. This fact has led a number of research groups to consider whether estimates of nutrient uptake or export based on satellite estimates of ocean color can be used to quantify the

magnitude of pathways of vertical exchange. This paper represents a new evaluation of whether particle export can be used to evaluate the realism of circulation and suggests that we can now obtain qualitative agreement between models and observational estimates, which in some cases can be used to suggest where there are deficiencies in models. Section 2 develops the concepts that are necessary to compare biological cycling predicted by diagnostic general circulation models with that inferred from satellites. Section 3 describes the observational data sets and models we use in this paper. Section 4 presents results of several model studies compared with different observational datasets. Implications of our results are discussed in Section 5.

2. Background

It is not straightforward to compare satellite-based estimates of biogeochemical cycling with those derived from circulation models because satellites and circulation models are sensitive to different processes. Somewhat simplistically, ocean color is directly related to *primary production*, the total uptake of some nutrient (e.g nitrogen or carbon) by phytoplankton. A key constant of proportionality is the optimal photosynthetic efficiency. This parameter differs significantly between different production schemes, so that primary production levels estimated by different investigators from the same chlorophyll levels can differ by a factor of two. Behrenfeld and Falkowski, (1997) provide a more detailed discussion of this issue. Of the total nutrient taken up, only some portion (referred to as *new production*) is supplied from nutrients which have not been previously cycled through organic material within the

euphotic zone. The remainder comes from compounds generated within the euphotic zone, which generally have a short lifetime. Inventories of living organic matter within the euphotic zone are generally quite small (Longhurst, 1998) indicating that most nutrients that are taken up by plankton are quickly consumed by grazers. Some fraction of this uptake is then turned into either dissolved or particulate detritral organic matter and some portion of this organic matter is exported from the surface layer. The total export of organic matter is referred to as *export production* and that portion associated with particulate matter is referred to as *particle export*.

The Ocean Carbon Model Intercomparison Project (OCMIP) developed a framework to predict biogeochemical cycling based on the work of Najjar et al. (1992). In this framework, circulation models are used to estimate the physical convergence of inorganic nutrients into the euphotic zone. The observed spatiotemporal distribution of inorganic nutrients can then be used to estimate the uptake (new production) and removal (export production) of these nutrients required to balance this convergence. The new and export production are not identical because dissolved organic matter can be produced at one location and transported laterally before being exported. Converting satellite-based estimates of primary production to new production requires a parameterization of the ratio of new to primary production (known as the *f-ratio*). Similarly, satellite-based observational estimates of particle export require a parameterization of the ratio of particle export to primary production. Dunne et al., (subm.) define this ratio as the *pe-ratio*. In a one-dimensional, time-mean model without dissolved organic matter, the f-ratio and pe-ratio will be the same. Laws et al. (2000) proposed a formulation of the f-ratio (which they assumed to be equal to the pe-ratio)

that depended on primary production and temperature and presented estimates of new production based on two estimates of primary productivity estimated using the Coastal Zone Color Scanner (CZCS) satellite. Dunne et al. (subm.) suggest a somewhat different formulation of the pe-ratio (*not* assumed to be equal to f-ratio) and use this formulation to estimate the particle export using newer data from the SeaWiFS satellite. This paper considers whether this new dataset can be used to give insight into the pathways of vertical exchange.

In a previous paper (Gnanadesikan et al. 2002), we compared the new production predicted by a suite of models with that given by CZCS ocean color and the f-ratio parameterization of Laws et al. (2000). The results were disappointing. Although the globally integrated new production bracketed the observational estimates, with the data favoring models with lower values of vertical mixing, the horizontal distribution of new production was very different in the models and the observations. This can be seen in Figure 1, which shows the zonally-integrated new production in one of the models and two observational estimates. The model shows three clear peaks, one in the southern ocean, one in the tropics and one in the high northern latitudes, with lower values in the subtropical gyres. In other words, where there is upwelling or deep wintertime convection, new production is predicted to be high. However, neither of the observational estimates shows a pattern with three peaks. The solid line without symbols (corresponding to an estimate of primary production made using an optimal photosynthetic efficiency given by an eighth-order polynomial in temperature as in Behrenfeld and Falkowski, 1997) shows no peak in the equatorial zone. Instead, this observational estimate is essentially constant from 30°S to 30°N. By contrast, the dashed

line (corresponding to an estimate of primary production made using an optimal photosynthetic efficiency which increases exponentially with temperature, as in Eppley, 1972) has no peak in the Southern Ocean but is essentially constant from 60°S to 20°S.

This basic picture held for all six models described in Gnanadesikan et al. (2002). As a result, when the spatial correlation was computed between the estimate based on the Behrenfeld and Falkowski (1997) algorithm (solid line in Figure 1) and the six models described in the paper, the results ranged between 0.14 and 0.21. When exponential photosynthetic efficiency was used instead, (dashed line in Figure 1) the correlations were higher, ranging between 0.39 and 0.46. This still means, however, that at best only 20% of the "observed" spatial variance in the data could be explained by the models.

A number of new developments make it worthwhile to revisit this problem. The first is that both SeaWiFS and reanalyzed CZCS data (Gregg et al., 2002) yield significantly higher values of primary productivity in the tropics than do the CZCS data that was available to us in Gnanadesikan et al. (2002). The second is that we have developed a new algorithm for estimating particle export ratio (Dunne et al., subm.), which matches the observed pe-ratio substantially better than the formulation of Laws et al. (2000) that we used in our previous paper. Finally, our group has developed new numerical models that agree better with deep oxygen and radiocarbon concentrations. We find that these new models and new observational estimates produce a picture of biological cycling that is in better agreement than was seen in Figure 1. We also demonstrate that particle export adds useful information about the ocean circulation which is not captured by the other tracers.

3. Methods

Circulation Model Description

We describe simulations from four different general circulation models. All four are coarse-resolution, four degree general circulation models based on the Modular Ocean Model, Version 3 code of Pacanowski and Griffies (1999). All are forced with monthly-varying fluxes of heat, salt and momentum, as well as diagnosed flux corrections of heat and salt computed by restoring the surface temperatures and salinities back towards the observations collected by Levitus and Boyer (1994). The monthly heat and salt fluxes for all the models (except for model PRINCE2A, as described below) are taken directly from the climatology of DaSilva et al. (1994). Additionally, a representation of the biological cycle of phosphate and carbon is included, as specified by the OCMIP protocols and described in Gnanadesikan et al. (2002). All models were run out to equilibrium. The differences between the models are summarized in tabular form in Table 1. The models are:

1. PRINCE1: This is the KVLOW+AILOW case from Gnanadesikan et al. (2002). This model has been found to have a good temperature, salinity and pycnocline structure, and reasonable globally integrated rates of uptake of anthropogenic CO_2 and CFCs (Matsumoto et al., in prep.). The PRINCE1 model was used as the initial submission to the OCMIP-II program. However, it has a number of serious deficiencies. As can be seen in Figure 2, the PRINCE1 model has far too little radiocarbon in the deep

Pacific. This lack of radiocarbon is mirrored by a lack of oxygen, which largely accounts for the low values of zonally averaged oxygen in Figure 2c. Observed values of oxygen in the data in the deep Pacific range between 150 and 200 μm/l, while values in the model are much lower, dropping below 75 μm /l over much of the abyssal ocean. Similarly, the observed radiocarbon in the deep north Pacific is greater than -250 permil in the zonal average, while it drops below -350 permil in the PRINCE1 model.

2. HIMIX: This is the KVHIGH+AIHIGH case from Gnanadesikan et al. (2002).The lateral mixing coefficients was increased from 1000 to 2000 m^2/s and the vertical mixing coefficient within the pycnocline was increased from 0.15 to 0.6 cm^2/s. These changes were made together so as to preserve the shape of the pycnocline while shifting the pathway of deep upwelling from the Southern Ocean to low latitudes, following the theory of Gnanadesikan (1999a). As discussed in Gnanadesikan et al. (2002) this case also has substantially more convection in the Southern Ocean. As a result the deep ocean is better ventilated. The major deficiency of this model is that the vertical mixing coefficient is far larger in the tropical pycnocline than can be supported by direct measurements (Ledwell et al., 1993). Insofar as HIMIX produces "good" simulations of radiocarbon and oxygen, it most likely does so for the wrong reasons.

3. PRINCE2: The base for this model was the KVHISOUTH+AILOW case of Gnanadesikan et al. (2002). Two major alterations were made to this model to try to

improve the deep radiocarbon concentrations in the North Pacific which was also too low in the KVHISOUTH+AILOW model. The first was to force the model to form deep water in the Ross and Weddell Seas by imposing a strong restoring to the salinities found in bottom water at four points during the three months of winter. It was found that in the Levitus and Boyer (1994) data set, towards which this model was restored, the deep water found at 200-400 m in the Weddell and Ross Seas never outcropped at the surface. Since it is well known that these waters do in fact outcrop during the winter in coastal polynyas (Foster and Carmack, 1976; Fahrbach et al., 1994) it seemed reasonable to force the formation of such water masses at a few points. This change was found to increase deep water concentrations of radiocarbon by 60 per mil, reducing the observed mismatch by about half. Since case HIMIX showed us that a large increase in vertical diffusion coefficient could increase deep-water radiocarbon concentrations, we decided to investigate the impacts of a small increase in vertical diffusivity. We thus increased the vertical diffusion coefficient in the pycnocline from 0.15 cm^2/s to 0.3 cm^2/s. This latter value is significantly larger than that found by Ledwell et al. (1993). However, it is identical to that used in previous versions of the POBM (Najjar et al., 1992, Gnanadesikan, 1999b). As seen in the top row of Figure 3, the zones of low oxygen and radiocarbon are much reduced in this model relative to PRINCE1 and the overall picture looks fairly plausible (a more quantitative comparison is made later in the paper).

4. PRINCE2A: As in PRINCE2, boundary conditions were altered in the far Southern Ocean to increase the formation of deep dense water. However, in contrast to

PRINCE2, this model attempts to resolve the radiocarbon problem without increasing the vertical diffusion coefficient in the pycnocline. Instead, a series of changes were made to make the Southern Ocean circulation more "realistic" (at least putatively). The wind stress field was changed from Hellerman and Rosenstein (1982) to the ECMWF reanalysis of Trenberth et al. (1989). The zonally-averaged differences in the wind stress are shown in Figure 4a. The Drake Passage, which was too broad in the PRINCE1, HIMIX and PRINCE2 setup (adopted from the old GFDL coupled model) was narrowed by one grid point (~450 km). Finally, analysis of the PRINCE1 model showed that the flux corrections calculated from restoring and the applied fluxes were operating in the opposite direction throughout much of the Southern Ocean (Figure 4b). This implies that the surface waters in PRINCE1 are too fresh. The applied fluxes are quite uncertain (they are not, for example globally balanced). We decided to adjust the applied fluxes by adding the flux corrections computed from restoring. As can be seen from Figure 4b, this substantially reduces the applied freshwater flux to the Southern Ocean. The result of this change is to destabilize the surface, allowing for more deep ventilation. This model also has distributions of oxygen and radiocarbon that are much more realistic (bottom row of Figure 3).

Data sets with which the models are compared

Comparisons are made between the models and three datasets, two of which are new. Oxygen is compared against the Conkright et al. (1998) World Ocean Atlas. Radiocarbon is compared against a gridded dataset compiled from WOCE and GEOSECS data by R.M. Key and available on the Pacific Marine Environmental

Laboratory Live Access Server. Additional comparisons are made with estimates of particle export made from satellite-based estimates of chlorophyll concentration. For these we use the empirical regression of Dunne et al. (subm.) which links particle export to primary production and chlorophyll using equation (1)

Particle export = Primary Productivity

$$*\max(0.042,\min(0.72,-0.0078\cdot T + 0.0806\cdot \ln(Chl) + 0.433)) . \quad (1)$$

This formulation was compared against an extensive compilation of observations and was found to account for 87% of the variability in particle export and 64% of the variability in pe-ratio. By contrast, the formulation of Laws et al. (2000) explains only 48% of the variability in pe-ratio. As noted by Dunne et al. (subm.), because there is significant uncertainty in the estimates of particle export, explaining 64% of the variance is about as good a fit as possible.

Significant uncertainties remain about how to estimate primary productivity from observed chlorophyll. We use three estimates. The first is the algorithm of Behrenfeld and Falkowski (1997) using an eighth-order polynomial estimate of the temperature-dependent optimal photosynthetic efficiency (P_B^{opt}) which has a maximum at around 20°C. The second estimate uses the algorithm of Carr (2002) which assumes an exponential temperature dependence for P_B^{opt} similar to that of Eppley (1972). The third estimate uses the primary productivity estimate of Lee et al. (1996). The algorithm for primary production used in this paper is documented in Marra et al. (in prep.). We use satellite estimates of chlorophyll concentration from SeaWIFS.

Figure 5 shows the estimated zonally averaged particle export using (1) for the three separate estimates of primary productivity. The results contrast significantly with the satellite-based estimates of new production in Figure 1. All three estimates of primary productivity result in predictions of particle export that have the three lobes predicted by the numerical models. This is the result of higher chlorophyll retrievals from both the tropics and at the Southern edge of the Southern subtropical gyre as well as the higher estimates of pe-ratio in the tropics. There are substantial differences between the different estimates. The Carr (2002) parameterization produces a value of export that is twice that of the BF97 algorithm in the tropics, while the Marra (in prep.) parameterization is higher in the tropics and substantially lower in the subpolar regions.

4. Results

The different models exhibit very different patterns of convection, and thus in the role that convection plays in ventilating the deep ocean. Figure 6 shows the zonally and annually averaged convective index (set to 1 if convection occurs and zero otherwise). A value of 1.0 means that convection is always occurring at every point at a particular latitude and depth. A value of 0.5 can mean that half the points convect all the time, that all the points convect half the time, or some other combination of spatial distribution and temporal frequency of convection. In PRINCE1, there is persistent convection at a latitude of 70°S and a depth of 1200 m. This convection carries cold Antarctic waters into the deep ocean in this model. Analysis of the T-S properties shows that convection in the model occurs near the point where cold fresh Antarctic waters become denser than warmer, salty Circumpolar deep waters. When surface conditions are changed

(PRINCE2) so as to make the cold Antarctic Waters in the Ross and Weddell Sea more salty, the amount of convection in the interior actually decreases. Increasing the vertical mixing in the tropics (HIMIX) results in more of a heat flow to the high latitudes and also shuts off this deep convective pathway. Both PRINCE2 and HIMIX have more convection in the upper 500 m, especially close to the Antarctic continent. The only case that shows a major increase in convection throughout the water column relative to PRINCE1 is PRINCE2A. Presumably this results from the changes to the fluxes reducing the surface fresh bias in the Southern Ocean.

All three new models have distributions of oxygen that are much more in line with observations than those produced by PRINCE1. Analysis of the oxygen errors (Figure 7) shows that the models get essentially the right concentrations below 1000m. However, both PRINCE2A and HIMIX have slightly too much oxygen below a few hundred meters. Analysis of the radiocarbon distributions in the Pacific (Figure 8) shows a very similar picture, though now both PRINCE2A and PRINCE2 get about the right answer. In contrast to the oxygen distributions, the radiocarbon distributions show a pattern of error that is more uniform in the north-south direction. In this paper, we focus on the oxygen distribution. Matusmoto et al. (in prep) discuss the radiocarbon distributions in more detail.

Although all three new models get more oxygen into the deep ocean than PRINCE1, they accomplish this in greatly different ways. An analysis of the term balances in the model (Table 2) shows that HIMIX and PRINCE2 largely accomplish the increased transport of oxygen to the abyssal ocean (below 1000 m) by increasing the diffusive transport to the deep ocean (primarily south of 30°S). By contrast, PRINCE2A

increases the deep transport by increasing the convective transport across 1000 m.

Examination of regional oxygen profiles (Figure 9) gives a more detailed picture of the differences between the models and the data. The observed oxygen profiles in both the Southern Ocean and North Pacific show mid-depth minima. These minima are pronounced in the North Pacific but clearly present in the Southern Ocean as well. All of the models have some representation of the North Pacific minimum (though it is too deep in all cases and not pronounced enough). None of the models, however, produce a mid-depth minimum in the Southern Ocean. In all of the models the lowest values of oxygen are found at the bottom

The failure of all the models to get a near-bottom oxygen maximum in the Southern Ocean is most likely related to the failure of level-coordinate general circulation models to simulate downslope flows associated with the formation of dense water in marginal seas. As described in Winton et al. (1998) coarse-resolution models fail to resolve the dynamics governing these flows for a number of reasons. Low vertical resolution means that the Ekman layer is not properly resolved. Low horizontal resolution, coupled with the use of level coordinates, means that dense flows entering the interior entrain far too much light water, diluting the surface signal. Resolution of this problem would require the use of a model (isopycnal or sigma) that is capable of resolving such flows.

In order to judge the ability of models to simulate observed patterns, it is helpful to examine multiple patterns simultaneously. Taylor (2001) proposed a diagrammatic way of comparing multiple patterns within and across models. The diagram he proposed is a polar plot in which the radius is the ratio of the standard deviation of the model to

that of the data, and the angle is the inverse cosine of the correlation. In other words, the radius shows whether the model gets the right amplitude of variability, while the angle shows whether the variability has the right spatial pattern. The distance from the (1,0) point on the plot is then the normalized unbiased RMS error. Figures 10a-c show Taylor diagrams for temperature, salinity, and oxygen for the three models compared to the World Ocean Atlas 1998 data (Conkright et al. 1998). Figure 10d shows radiocarbon in per mil from the three models compared against the gridded data set for the Pacific. As can be clearly seen, the PRINCE2 and PRINCE2A runs have slightly worse solutions for temperature and salinity (though they do explain more than 90% of the observed variance in both cases) but much better solutions for oxygen and radiocarbon. In general, oxygen is the worst-performing tracer, with less than 80% of the observed variance explained in all of the runs.

What does particle export tell us about the four models? Figure 11 shows a comparison between the four models and the three observationally based estimates using SeaWiFS chlorophyll and the three different primary productivity algorithms. Basin-scale integrals are presented in Table 3. By comparing Figure 11 with Figure 1, we can see that the qualitative agreement between models and observational estimates is much better than in Gnanadesikan et al., (2002), with all the estimates of particle export showing three latitudinal maxima. The Southern Hemisphere minimum in productivity is shifted to the south in all of the models.

Comparing Figures 1 and 11 also highlights some interesting differences between Gnanadesikan et al. (2002) and this study. In Gnanadesikan et al. (2002) the tropics were the region with the biggest differences between the models and the

observational estimates while at least one model was consistent with the observational estimates between 50°N and 70°N (Figure 16 of Gnanadesikan et al., 2002). In the current study, the observational range in the tropics (3.2-5.0 GtC/yr) brackets the two models with low vertical mixing (PRINCE1 and PRINCE2A). The model with an intermediate level of mixing (PRINCE2) has a tropical particle export of 5.6 Gt/yr, slightly higher than the largest observational estimate (that of Carr, 2002). The HIMIX model has a tropical particle export of 7.7 GtC particle export, 50% higher than the highest of the three observational estimates and more than 140% larger than the Behrenfeld and Falkowski (1997) observational estimate. The argument made in previous papers that high levels of diapycnal diffusivity are inconsistent with observed tropical productivity (Gnanadesikan and Toggweiler, 1999; Gnanadesikan et al., 2002) is supported by the overprediction of particle export by HIMIX.

In contrast to Gnanadesikan et al. (2002), where the biggest area of disagreement was the tropics, the biggest area of disagreement is now the Northern Hemisphere. As discussed in Gnanadesikan et al. (2002), a major source of disagreement between the models and the observational estimates is the lack of export production in the North Pacific. The North Pacific particle export in the models ranges between 0.29 and 0.62 GtC/yr. This is much smaller than the estimates of particle export from satellite color, which range between 1.3 and 1.8 GtC/yr. One signature of this failure is the bias towards high oxygen in the northern intermediate waters seen in Figures 7 and 9. At 800 m, oxygen concentrations in the North Pacific are higher than observations by up to 120 micromole/kg. The failure of the coarse-resolution models to simulate the North Pacific well was discussed in detail in Gnanadesikan et al. (2002). There it was argued that the

boundary conditions applied to the models failed to simulate the formation of a water mass with high preformed nutrients in the Northwest Pacific. The dynamics of how this water mass is formed remains a topic of current research.

A composite Taylor diagram, showing comparisons between the 4 hydrographic parameters and the monthly-varying, zonally-averaged particle export is shown in Figure 12. A zonal average was used to eliminate spurious low correlations arising from extreme values of particle export in coastal points not resolved by the model. The results show that the spatiotemporal distribution of particle export remains much more poorly simulated than any of the hydrographic parameters. Correlations between the observational estimates and the models range between 0.29 and 0.65 (r^2 of 0.1 and 0.4). At their best, the models explain 40% of the observed zonally-averaged spatiotemporal variance of export flux, substantially better than the upper limit of 20% found in Gnanadesikan et al. (2002) but still far below the model performance for radiocarbon or even oxygen. The correlations are substantially better for the Carr and Marra parameterizations than they are for the Behrenfeld and Falkowski parameterization, consistent with our earlier results. The results also show that although the distribution of oxygen and radiocarbon are significantly improved in the PRINCE2 and PRINCE2A models relative to PRINCE1, the new models do not capture a higher fraction of the variability in the export flux. Moreover, the distributions of temperature and salinity (the best-known hydrographic fields) are best simulated in PRINCE1!

5.Conclusions

Physical oceanographers seek to simulate both the density structure of the ocean and the pathways of vertical exchange by which this structure is maintained. This paper reinforces the point made in previous work that estimates of biogeochemical cycling can be very helpful in reaching this goal. We find that general circulation models which simulate temperature and salinity comparably well, can differ significantly in how they simulate of oxygen, radiocarbon and particle export. PRINCE2, PRINCE2A and HIMIX all have essentially comparable errors to PRINCE1 with regard to temperature and salinity, but are much better with regard to radiocarbon and oxygen. However, these models all supply radiocarbon and oxygen to the deep ocean in different ways - pointing out that even these fields do not provide an unambiguous way of constraining how the ocean mixes vertically.

Particle export adds an additional piece of information that is not necessarily captured by radiocarbon or oxygen. In contrast to measurements of tracer distributions, particle export provides a strong constraint on the *rate* at which the surface ocean is renewed. PRINCE2 and PRINCE2A have very similar distributions of radiocarbon, but regional values of particle export that differ by up to 50%. HIMIX has twice the tropical particle export of PRINCE2A. Estimates of the rate of biological processes such a particle export or new production thus have the potential to be very sensitive tools in judging between some aspects of model solutions.

In order for such evaluations to have value however, the data with which they are being compared must be of relatively high accuracy. In Gnanadesikan et al. (2002) we

found the patterns produced by the general circulation model (which had three peaks corresponding to regions where deep waters upwelled) did not agree with the satellite-based estimates (which tended to have only two peaks). This result suggested that either the models or the observational estimates were systematically in error. A major conclusion of this paper is that the observational estimates were the source of the discrepancy. Newer estimates with better data and more refined algorithms show a much better qualitative agreement between models and data. This fact gives us more confidence about using the quantitative estimates of particle export (uncertain though they are) to draw conclusions about pathways of vertical exchange.

We find that it is difficult to have a coarse resolution model that does everything well. Changes made to produce better deep ventilation (and so to improve radiocarbon and oxygen) tend to have a negative effect on some aspect of the simulation (salinity, temperature, or particle export). Increasing the vertical mixing coefficient in the tropics tends to cause the tropical particle export to exceed the observational estimates (in the case of HIMIX by a factor of 2-5). Changes in the fluxes in the Southern Ocean (PRINCE2A) have a tendency to shift the zone of maximum productivity further to the south (it is in roughly the right location in the other three models).

Is this inconsistency a more general problem? A similar mismatch between model-derived export production and satellite estimates of production was recently found by Schlitzer (2002) using an inverse model. Schlitzer's model has fairly vigorous vertical exchange in the Southern Ocean and does a reasonable job at simulating the deep radiocarbon in the North Pacific (Matsumoto et al., in prep.). However, it predicts export fluxes in the Southern Ocean far larger than can be supported given satellite-

based estimates of productivity. Inconsistency between ventilating the deep ocean through increasing Southern Ocean vertical exchange and uptake of trace gasses is also found by Matsumoto et al. (in prep.) who examine the uptake of CFCs and anthropogenic CO_2 in the OCMIP models as well as in the models described here. They find that the PRINCE2A model has too much uptake of CFCs and anthropogenic CO_2 in the Southern Ocean.

Taken together, these results suggest that there are significant problems with ventilating the deep ocean in large-scale, level-coordinate general circulation models. There are a few ways to deal with the inconsistency between the apparently low vertical exchange suggested by the temperature, salinity, and tropical particle export, and the high vertical exchange suggested by the radiocarbon and oxygen. The first is that the ocean may not be in steady state, and so the radiocarbon and oxygen in the deep ocean at present may have been injected there at some time in the past when ventilation was more vigorous. This suggestion was advanced by Broecker et al. (2001), but the data on which it was based was recently challenged by Orsi et al. (2002). A second possibility is that the problem is due to the fact that all the models used here are level coordinate models. Such models do not accurately represent the ventilation of the deep ocean by plumes of dense water spilling out of marginal seas. Insofar as level-coordinate models cannot resolve the pressure gradients driving such flows nor maintain the density contrasts driving the flows, they will thus tend to underestimate the advective ventilation of the deep ocean. In contrast to convection (which requires the downward movement of young waters to occur at the same location as the upward movement of old, nutrient rich waters) such advective pathways allow the two to be separated in space, so that

increasing ventilation of the Southern Ocean does not necessarily require increased export production there. We are currently exploring this hypothesis using an isopycnal layer model, which is known to represent such flows more accurately (Winton et al., 1998). Preliminary results show that such models are in fact capable of maintaining plumes of young bottom water over thousands of kilometers and tens of years, in contrast to comparably forced level coordinate models which tend to dilute the overflows and limit their depth of penetration.

Acknowledgments: This work was supported by the Department of Energy Office of Biological and Environmental Research under the Ocean Carbon Sequestration Initiative (DOE Grant DE-FG0200ER63009), the Carbon Modeling Consortium (NOAA Grant NA56GP0439) and the NOAA Geophysical Fluid Dynamics Laboratory. We thank Joellen Russell and Bryan Arbic for useful suggestions on an earlier draft of this paper and the SeaWIFS project (Code 970.2) and Goddard Earth Sciences Data and Information Services Center/Distributed Active Archive (Code 902) at the Goddard Space Flight Center, Greenbelt, MD 20771, for the production and distribution of these data, respectively. These activities are sponsored by NASA's Earth Science Enterprise.

Bibliography

Behrenfeld, M.J. and P.G. Falkowski (1997) Photosynthetic rates derived from satellite-based chlorophyll concentration, *Limnol. Oceanogr.*, **42**, 1-20.

Broecker, W.S., S. Sutherland and T.H. Peng (1999) A possible slowdown of Southern

Ocean deep water formation, *Science*, **286**, 1132-1135.

Bryan, F. O. (1987) Parameter sensitivity studies of a primitive-equation ocean general circulation model, *J. Phys. Oceanogr.*, **17**, 970-985.

Carr, M.E. (2002) Estimation of potential productivity in Eastern Boundary Currents using remote sensing, *Deep Sea Res. II*, **49**, 59-80.

Conkright,M., S. Levitus, T. O'Brien, T. Boyer, J. Antonov, and C. Stephens, (1998) *World Ocean Atlas 1998 CD-ROM Data Set Documentation.*. Tech. Rep. **15**, NODC Internal Report, Silver Spring, MD, 16pp.

Da Silva, A., A.C. Young, and S. Levitus (1994) Atlas of Surface Marine Data 1994, Volume 1: Algorithms and Procedures, NOAA Atlas NESDIS 6, U.S. Department of Commerce, Washington, D.C.

Dunne, J.P., R.A. Armstrong, C.A. Deutsch, A. Gnanadesikan, J.L. Sarmiento and P.S. Swathi, Controls on global particle export and remineralization: Model development and calibration, subm. *Global Biogeochemical Cycles.*

Eppley, R.W. (1972) Temperature and phytoplankton growth in the sea, *Fisheries Bull.*, 70, 1063-1085.

Fahrbach, E., R.G. Peterson, G. Rohardt, P. Schlosser and R. Bayer (1994) Suppression of bottom water formation in the Southeastern Weddell Sea, *Deep Sea Res.* I, 41, 389-411.

Foster, T.D. and E.C. Carmack (1976) Frontal zone mixing and Antarctic Bottom Water formation in the Southern Weddell Sea, *Deep Sea Res.*, 23, 301-317.

Gent, P., and J.C. McWilliams (1990), Isopycnal mixing in ocean circulation models, *J. Phys. Oceanogr*., 20, 150-155.

Gnanadesikan, A. (1999a) A simple model for the structure of the oceanic pycnocline, *Science*. **283**, 2077-2079.

Gnanadesikan, A., (1999b) A global model of silicon cycling: Sensitivity to eddy parameterization and dissolution, *Global Biogeochemical Cycles*, 13, 199-220.

Gnanadesikan, A., R.D. Slater, N. Gruber and J.L. Sarmiento (2002) Oceanic vertical exchange and new production: A comparison between models and observations, *Deep- Sea Res. II*, **49**, 363-401.

Gnanadesikan, A. and J.R. Toggweiler (1999) Constraints placed by silicon cycling on vertical exchange in general circulation models, *Geophysical Res. Lett.*, **26**, 1865-1868.

Gregg, W.W., M.E. Conkright, J.E. O'Reilly, F.S. Patt, M.H.H. Wang, J.A. Yoder and N.W. Casey (2002) NOAA-NASA coastal zone color scanner reanalysis effort, Applied Optics, **41**, 1615-1628.

Hellerman, S. and M. Rosenstein, (1983) Normal monthly wind stress over the World Ocean with error estimates, *J. Phys. Oceanogr.*, **13**, 1093-1104.

Klinger, B.A., S. Drijfhout, J. Marotzke and J.R. Scott (2003) Sensitivity of basinwide meridional overturning to diapycnal diffusion and remote wind forcing in an idealized Atlantic-Southern Ocean geometry, *J. Phys. Oceanogr.*, **33,** 249-266.

Laws, E.A., P.G. Falkowski, W.O. Smith, H. Ducklow and J.J. McCarthy (2000) Temperature effects on export production in the open ocean, *Global Biogeochem. Cycles*, **14**, 1231-1246.

Ledwell, J., A. Watson, and C.S. Law (1993) Evidence for slow mixing across the the pycnocline from an open-ocean tracer release experiment, *Nature*, **364**, 701-703.

Lee, Z.P., K.L. Carder, J. Marra, R.G. Steward and M.J. Perry (1996) Estimating primary production at depth from remote sensing, *Applied Optics*, 35, 463-474

Levitus S., R. Burgett and T.P. Boyer. (1994). World Ocean Atlas 1994 Volume 3: Salinity. NOAA Atlas NESDIS 3. U.S. Department of Commerce, Washington,

D.C. 99 pp.

Longhurst, A. (1998) Ecological geography of the sea, Academic Press, 398 pp.

Marra, J., C. Ho and C. Trees, (in prep.) An algorithm for the calculation of primary productivity from remote sensing data, LDEO Technical Report, Lamont-Doherty Earth Observatory, Palisades, NY.

Najjar, R.G., J.L. Sarmiento and J.R. Toggweiler (1992), Downward transport and the fate of organic matter in the ocean: Simulations with a general circulation model, *Global Biogeochem. Cyc.,* 6, 45-76.

Orsi, A.H., W.M Smethie and J.L. Bullister (2002) On the total input of Antarctic waters to the deep ocean: A preliminary estimate from chlorofluorocarbon measurements, *J. Geophys. Res.,* **107**, 3122.

Pacanowski, R.C., and S.M. Griffies, (1999) The MOM3 Manual, Alpha version, NOAA/ Geophysical Fluid Dynamics Laboratory, 580pp.

Schlitzer, R., (2002) Carbon export fluxes in the Southern Ocean: results from inverse modeling and comparison with satellite-based estimates, *Deep Sea Res. II,* 49, 1623-1644.

Taylor, K.E., (2001) Summarizing multiple aspects of model performance in a single

diagram, *J. Geophys. Res.*, 106, 7183-7192,2001.

Toggweiler, J.R., A. Gnanadesikan, S. Carson, R.J. Murnane and J.L. Sarmiento, (2002) The strength of carbon pumps in box models, GCMs and the real ocean, Part I: The solubility pump, in press, *Global Biogeochem. Cycles*.

Trenberth,K.E., J. Olson, and W. Large (1989) A Global Ocean Wind Stress Climatology based on ECMWF Analyses. Tech. Rep. NCAR/TN-338+STR, National Center for Atmospheric Research, Boulder, Colorado.

Winton, M.E., R.W. Hallberg and A. Gnanadesikan, Simulation of density- driven downslope flow in z-coordinate ocean models, *J. Phys. Oceanogr.* , *28*, 2163-2174, 1998

Worthington, L.V., The case for near-zero production of Antarctic Bottom Water, *Geochim. Cosmochim. Acta*, 41, 1001-1006, 1977.

	PRINCE1	PRINCE2	PRINCE2A	HIMIX
Pycnocline diffusivity (tropics)	0.15	**0.3**	0.15	**0.6**
Pycnocline diffusivity (Southern Ocean)	0.15	**1.0**	**1.0**	**0.6**
Enhanced diffusivity in top 50m	No	No	**Yes**	No
Lateral diffusivity	1000	1000	1000	**2000**
Applied Surface Fluxes	Da Silva et al. (1994)	Da Silva et al. (1994)	DaSilva et al (1994) + **Correction globally**	DaSilva et al. (1994)
Temperature and salinity towards which restoring occurs	Levitus and Boyer (1994)	Levitus and Boyer (1994) + **Correction near Antarctica in winter**	Levitus and Boyer (1994) + **Correction near Antarctica in winter**	Levitus and Boyer (1994)
Wind stress	Hellermann	Hellermann	**ECMWF**	Hellermann
Topography	Wide Drake Passage	Wide Drake Passage	**Narrow Drake Passage**	Narrow Drake Passage

Table 1: Tabular description of the four models. Key differences from PRINCE1 are shown in bold.

Process Experiment	Consumed	Advection	Convection	Total Vertical Diffusion	Total Vertical Diffusion (South of 30S)
PRINCE1	-3.05	3.01	0.03	0.03	0.09
PRINCE2	-4.51	3.03	0.22	1.26	1.35
PRINCE2A	-4.36	2.44	1.41	0.46	0.73
HIMIX	-5.92	3.32	0.43	2.17	2.27

Table 2: Oxygen demand and supply below 1000m in different models. Note that in PRINCE2 and HIMIX the enhanced demand (relative to PRINCE1) below 1000m is largely supplied by diffusion in Southern Ocean latitudes, while in PRINCE2A it is largely supplied by convection. The supply by vertical diffusion to the abyssal ocean is larger when only the Southern Ocean is considered because in low latitudes there is upward diffusion of oxygen into the oxygen minimum zone.

Region Estimate	Global	Northern Pacific (>30N)	Northern Atlantic (30N-75N)	Tropics (23S-23N)	Southern Ocean (<30S)
Behrenfeld Falkowski	11.1	1.8	1.7	3.2	3.4
Carr	12.0	1.6	1.4	5.0	2.9
Marra	9.7	1.3	1.2	3.9	2.4
PRINCE1	6.9	0.31	0.64	**3.9**	1.6
PRINCE2	**10.0**	0.39	0.85	5.6	**2.4**
PRINCE2A	8.8	0.29	0.79	**3.7**	**3.5**
HIMIX	13.0	0.62	0.95	7.7	**2.7**

Table 3: Export flux in GtC/yr from different regions using three different observational estimates of export and the four general circulation models described in the text. Model results which lie within the range of the observational estimates are highlighted in bold text.

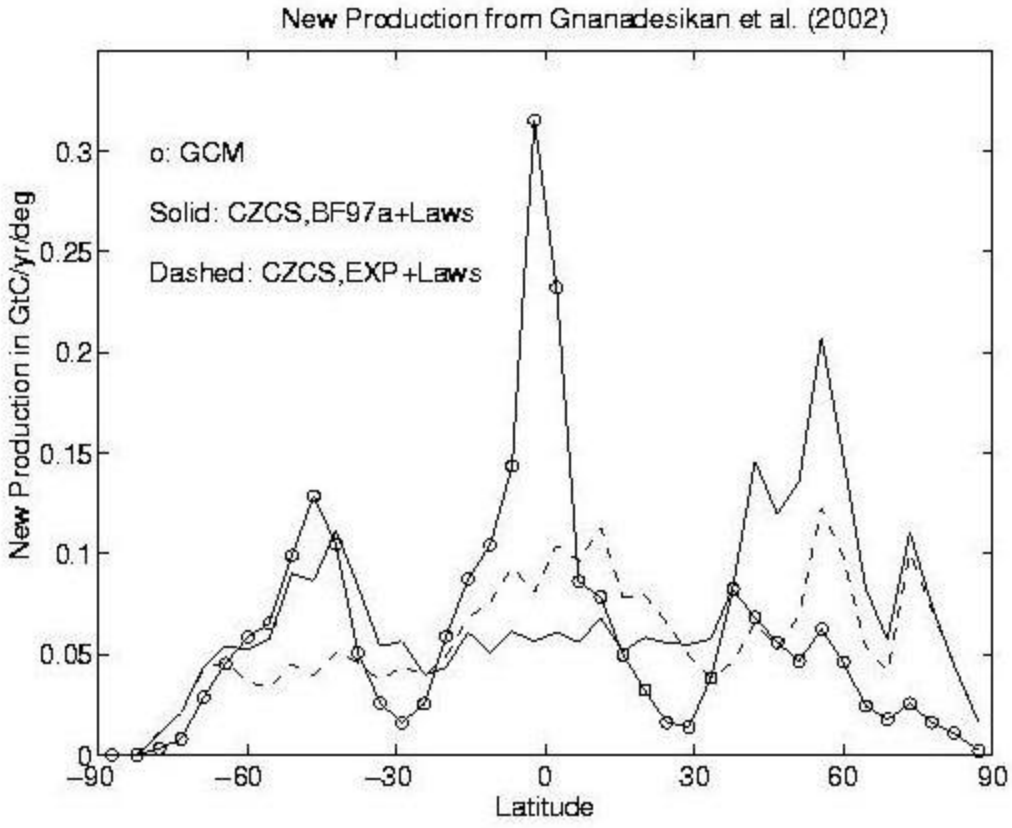


Figure 1: New production estimated from models and observations by Gnanadesikan et al., (2002). The two observational estimates use CZCS chlorophyll and the parameterization of f-ratio of Laws et al. (2000). Note that neither of the observational estimates exhibits three clear peaks.

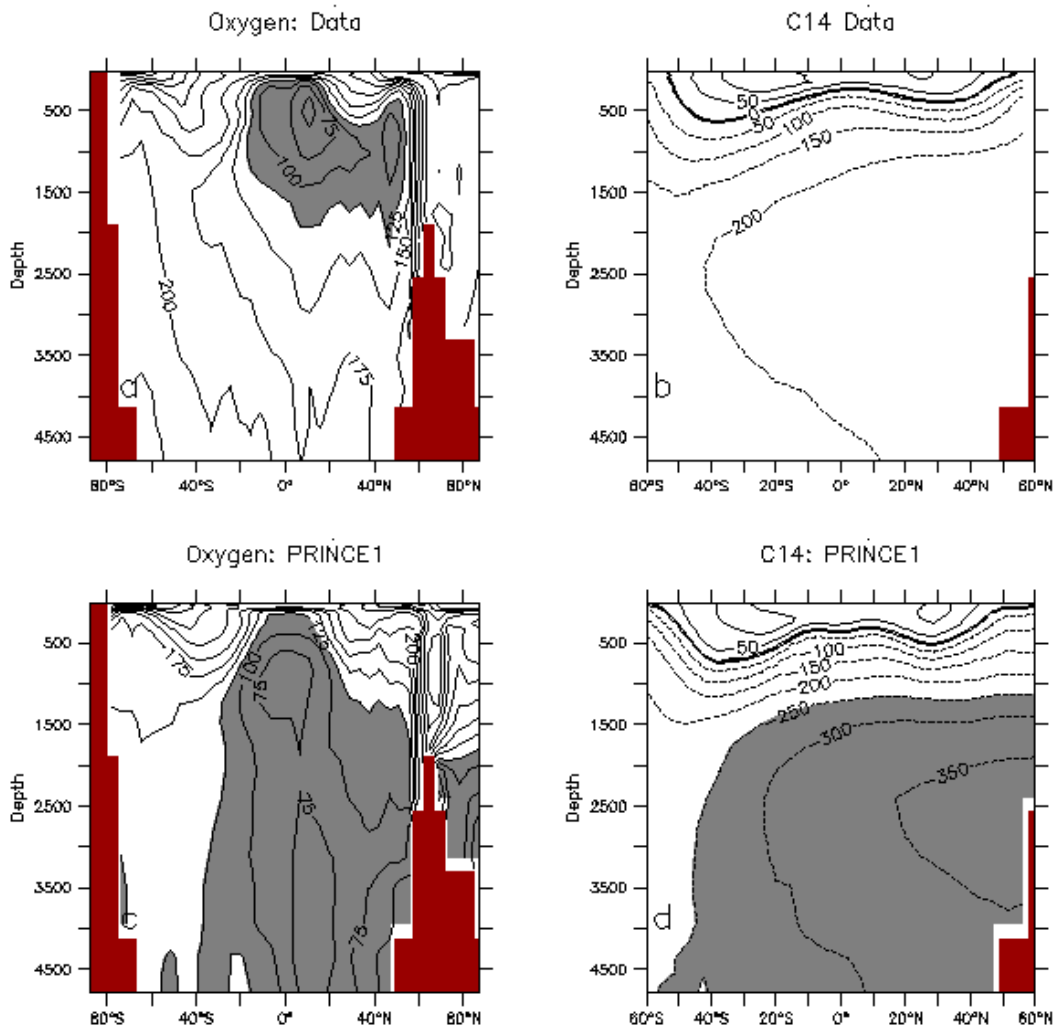


Figure 2: Illustration of poor ventilation of the deep ocean in the OCMIP-II PRINCE1 model. (a) Zonally-averaged oxygen from the Levitus (1994) atlas. Shading shows concentrations less than 125 µmol/l. (b) Zonally-averaged radiocarbon from the WOCE dataset in the Pacific. (c) Zonally-averaged oxygen from PRINCE1 biotic simulation. (d) Zonally-averaged radiocarbon in the Pacific from PRINCE1 historical simulation (year shown is 1995). Shading shows concentrations less than –250 permil, not seen in zonally-averaged observations.

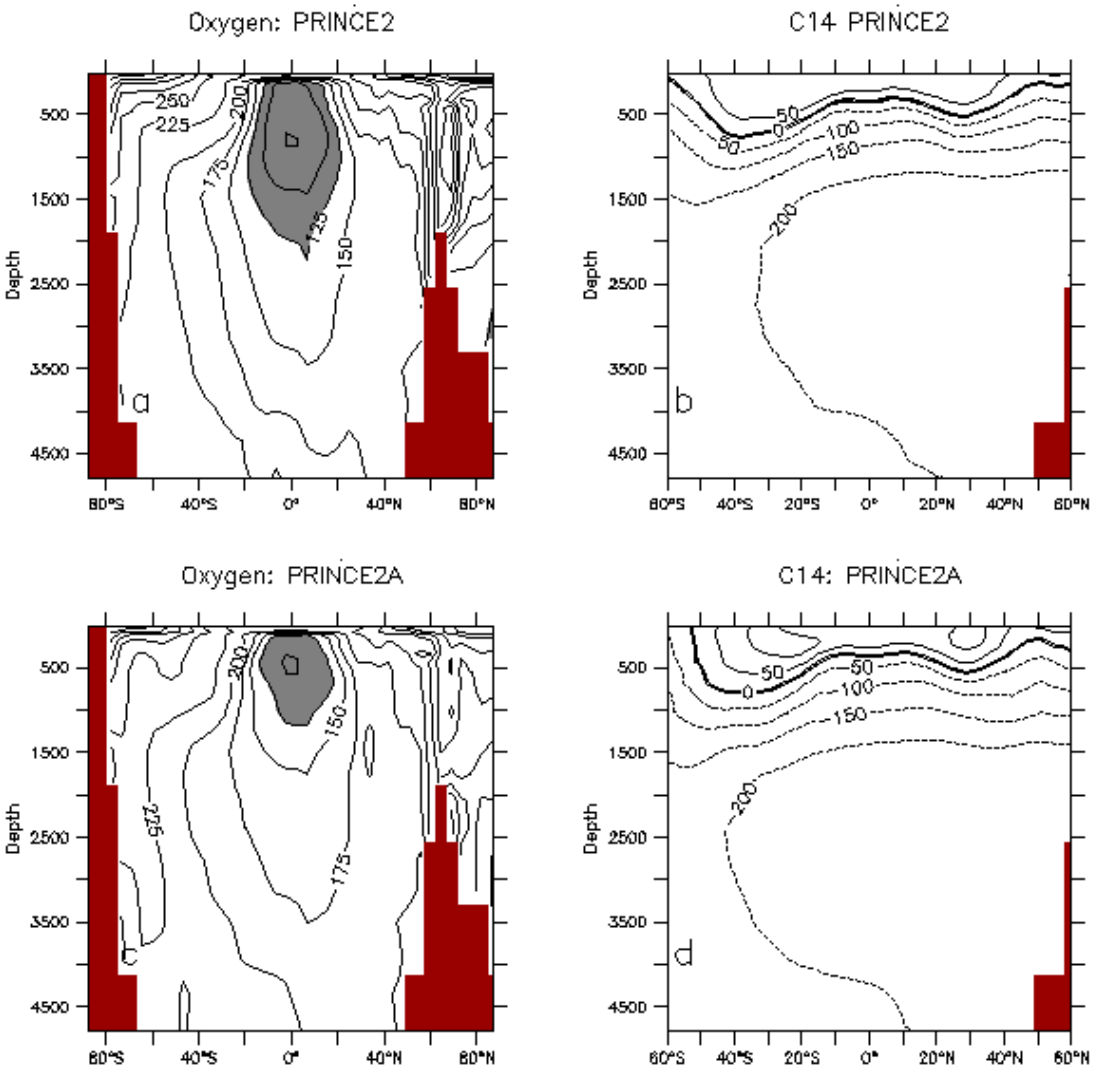
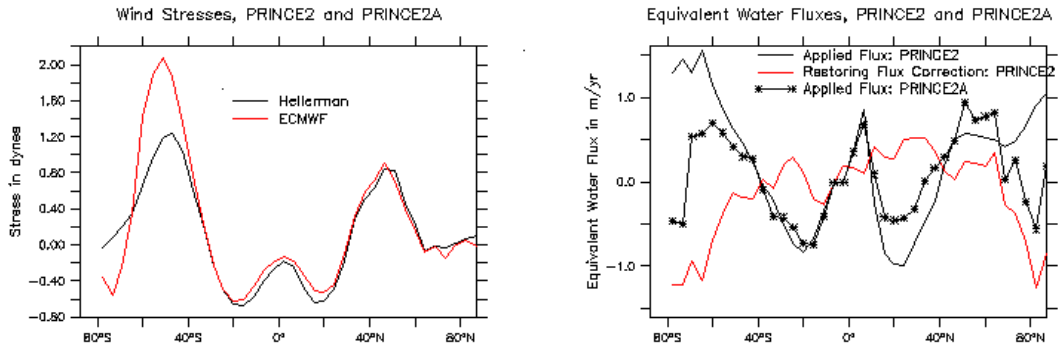


Figure 3: Same as Figure 2 but for PRINCE2 and PRINCE2A models. Note that the disagreement with the data is significantly reduced, with deep oxygen and radiocarbon concentrations increasing.



.

Figure 4: Changes in surface fluxes between PRINCE1 and PRINCE2A. (a) Wind stresses. Solid line is from Hellerman and Rosenstein (1982). Dashed line is from Trenberth et al (1989). Note that the ECMWF stresses are much higher winds in the Southern Ocean. (b) Applied (solid) and restoring (dashed) water flux, zonally averaged, for PRINCE1. Note that the fluxes are in opposite directions in the Southern Ocean. Since the deep ocean is salty, the implication is that the applied fluxes are making the surface too fresh

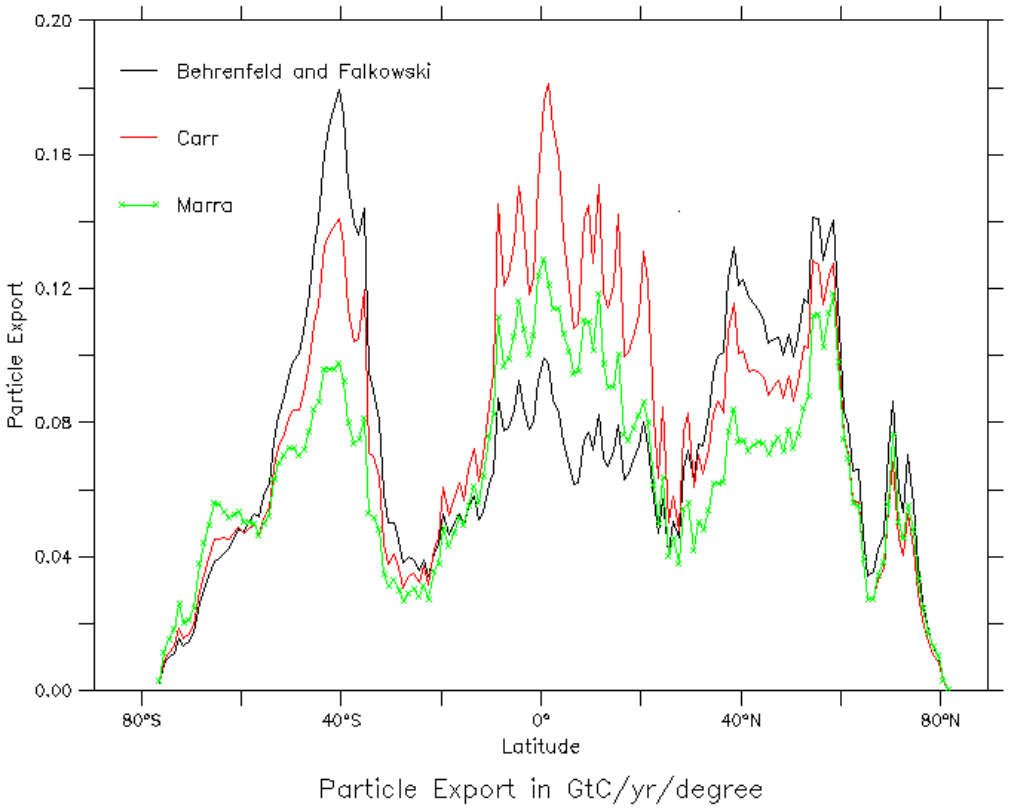


Figure 5: Particle export in GtC/yr/degree using three observational estimate of primary productivity (as outlined in the text) and the regression of Dunne et al. (in prep.). Note that there are three lobes rather than 2 as in Figure 1. This is due both the to use of new observational estimates for chlorophyll (which have more chlorophyll in the tropics) and the new pe-ratio parameterization which allows for higher pe-ratios in productivity tropical regions.

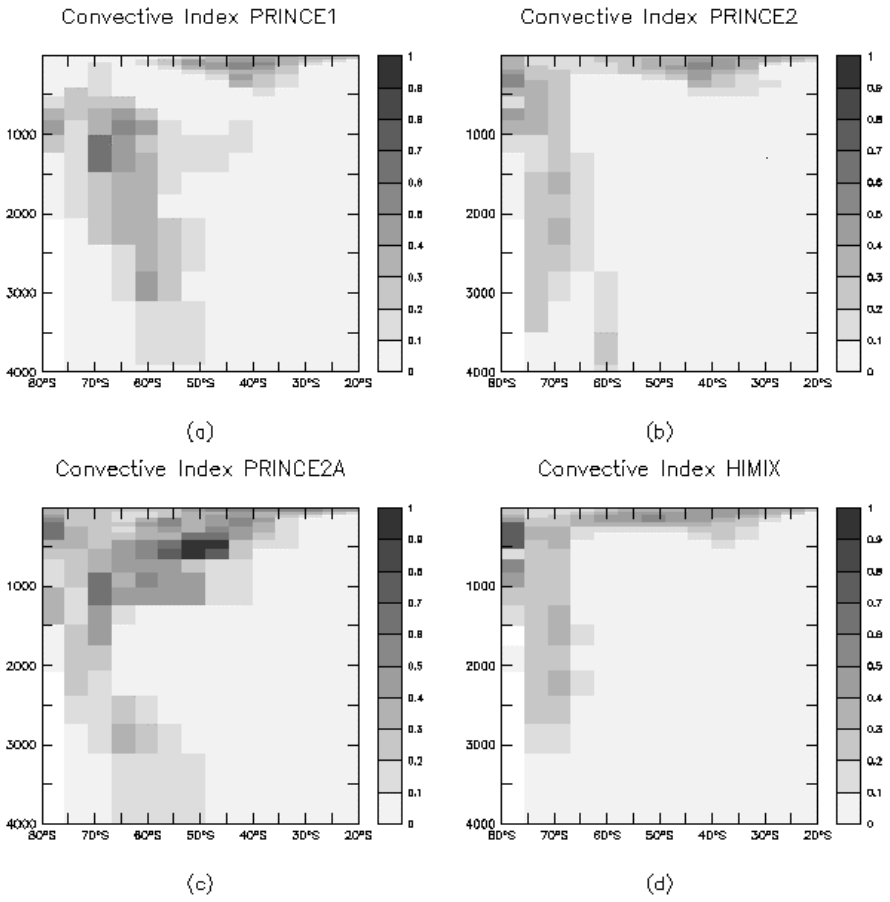


Figure 6: Convective index in the four models, averaged zonally and annually south of 20S. (a) PRINCE1. Note the existence of a subsurface zone of convection that does not connect to the surface. (b) PRINCE2. Convection is shifted closer to the continent. (c) PRINCE2A. Note the increase in convection between 500 and 1000m. (d) HIMIX.

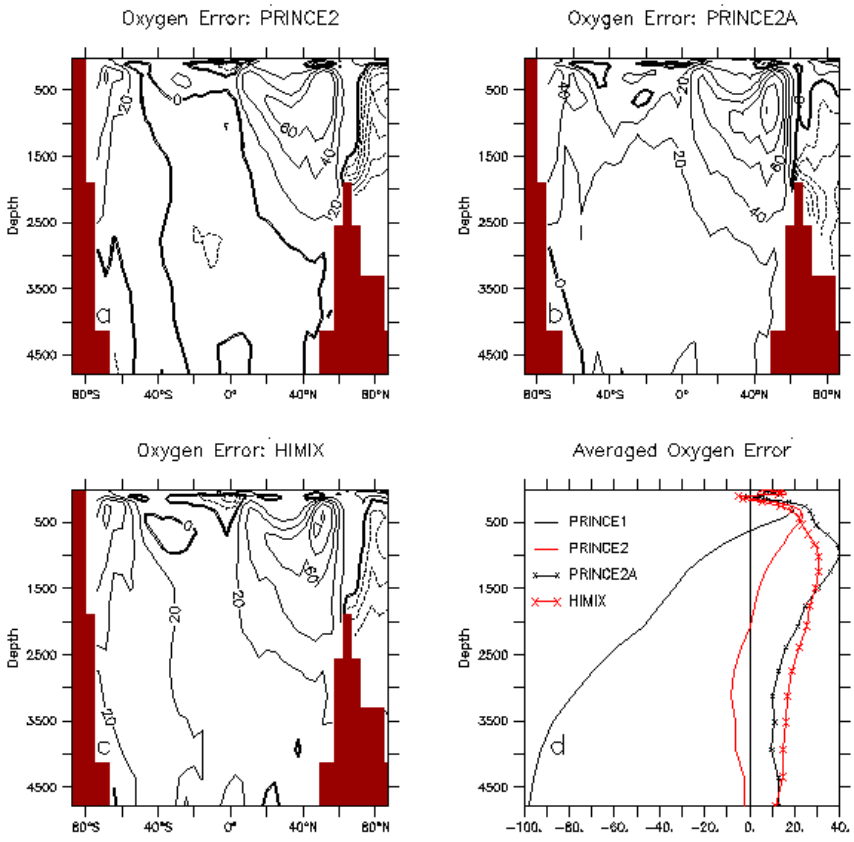


Figure 7: Oxygen errors, demonstrating that the changes to model physics and forcing do improve the ventilation of the deep ocean. Note, however that all the simulations have too much oxygen into the northern subpolar thermocline. (a) PRINCE2. (b) PRINCE2A. (c) HIMIX. (d) Horizontally averaged oxygen error including PRINCE1.

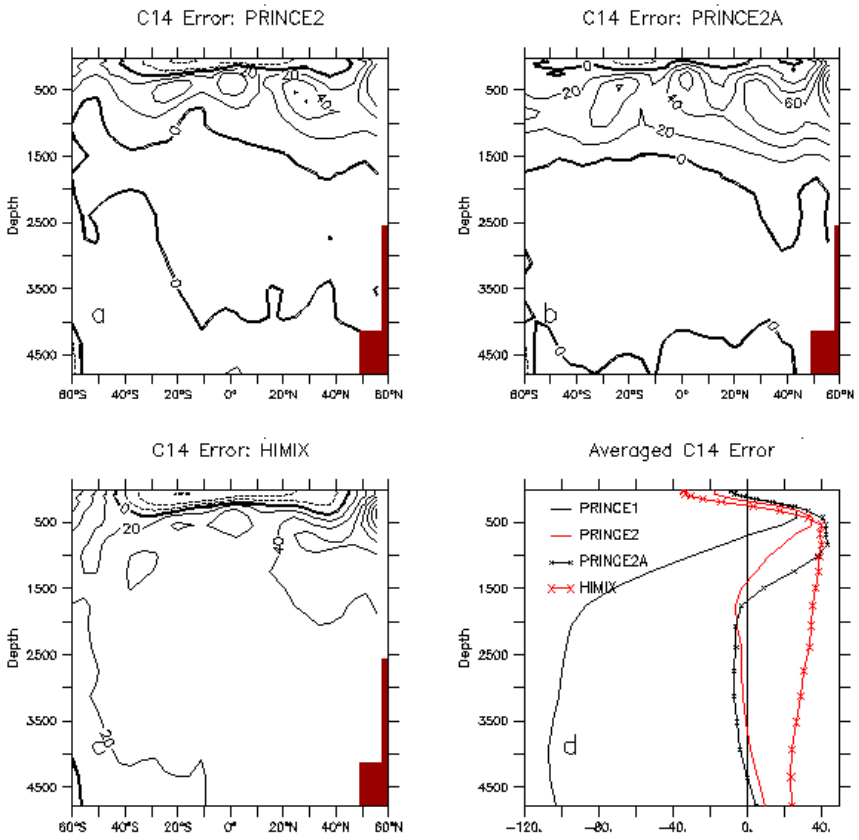


Figure 8: Radiocarbon errors from new runs.(a) PRINCE2. (b) PRINCE2A. (c) HIMIX. (d) Horizontally averaged radiocarbon error including PRINCE1.

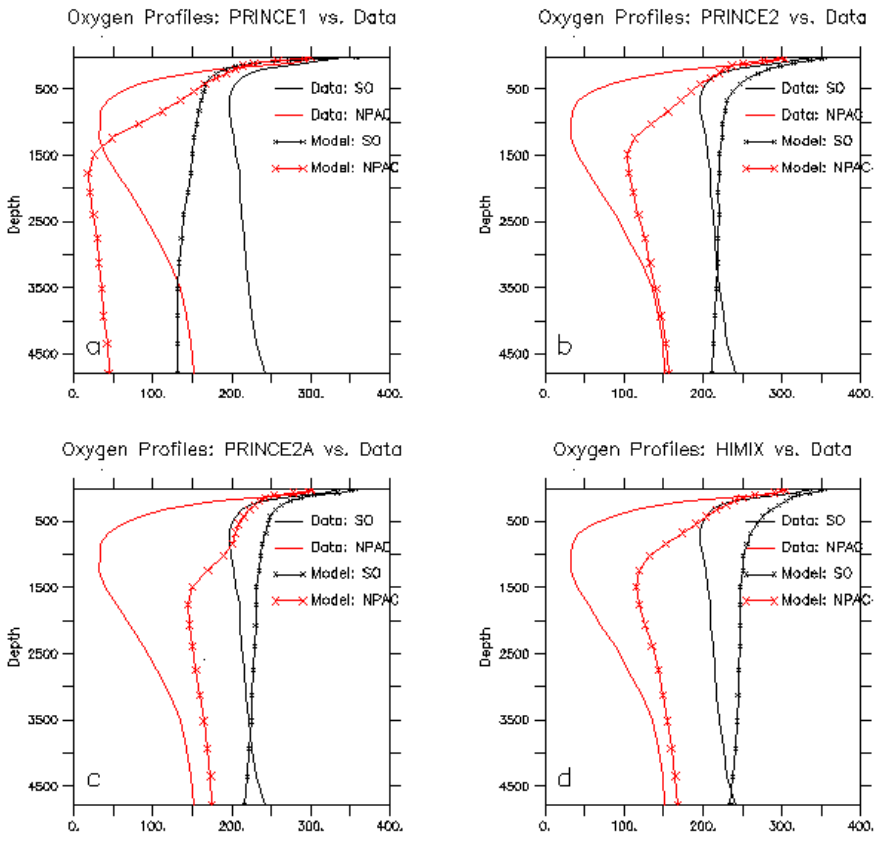


Figure 9: Oxygen profiles in the Southern Ocean and North Pacific in the four models. Note that the data has a mid-depth minimum, so that concentrations in the deep Southern Ocean are higher than those at 1000m. This signal is not seen in any of the models. Southern Ocean profiles are zonally averaged south of 60S. North Pacific profiles are between 40N and 60N. (a) PRINCE1. (b) PRINCE2 (c) PRINCE2A (d) HIMIX.

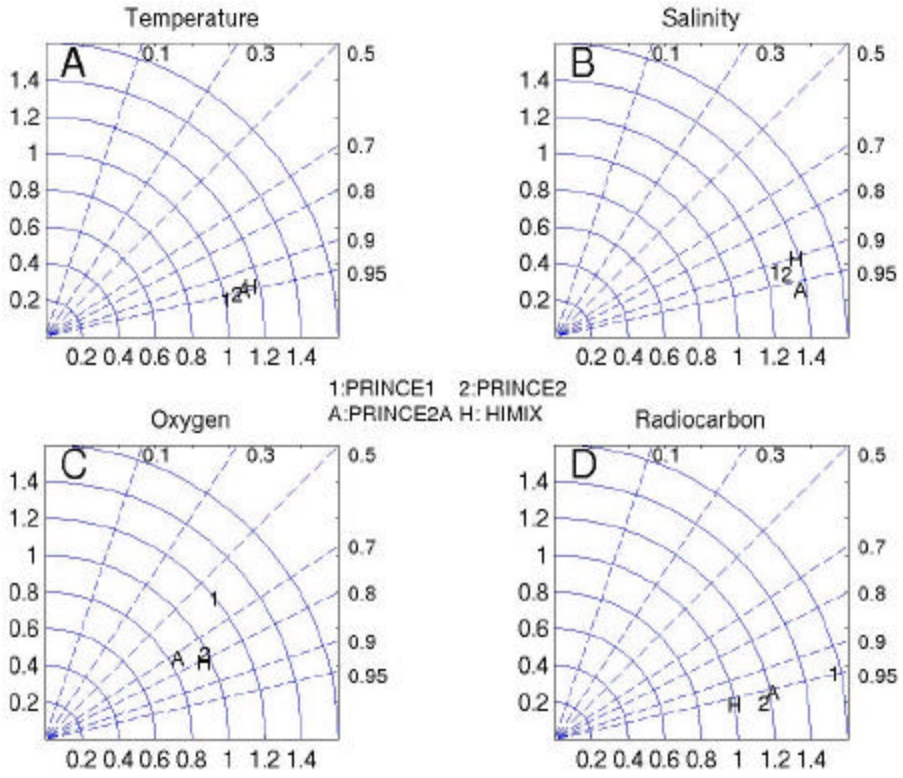


Figure 10: Taylor diagrams of the different model runs. Angles are labeled with r^2 values, showing fraction of variance explained. The angle thus provides a measure of how well the observed 3-dimensional pattern of variability is reproduced. The radius is the scaled range, providing a measure for how well the amplitude of the variability is reproduced. The distance from the (1,0) point is the scaled RMS error. (a) Temperature. (b) Salinity. (c) Oxygen. (d) Radiocarbon. Note that the best simulation for radiocarbon (H) has the worst simulation for temperature and salinity while the best simulation for temperature (1) has the worst simulation for radiocarbon and oxygen.

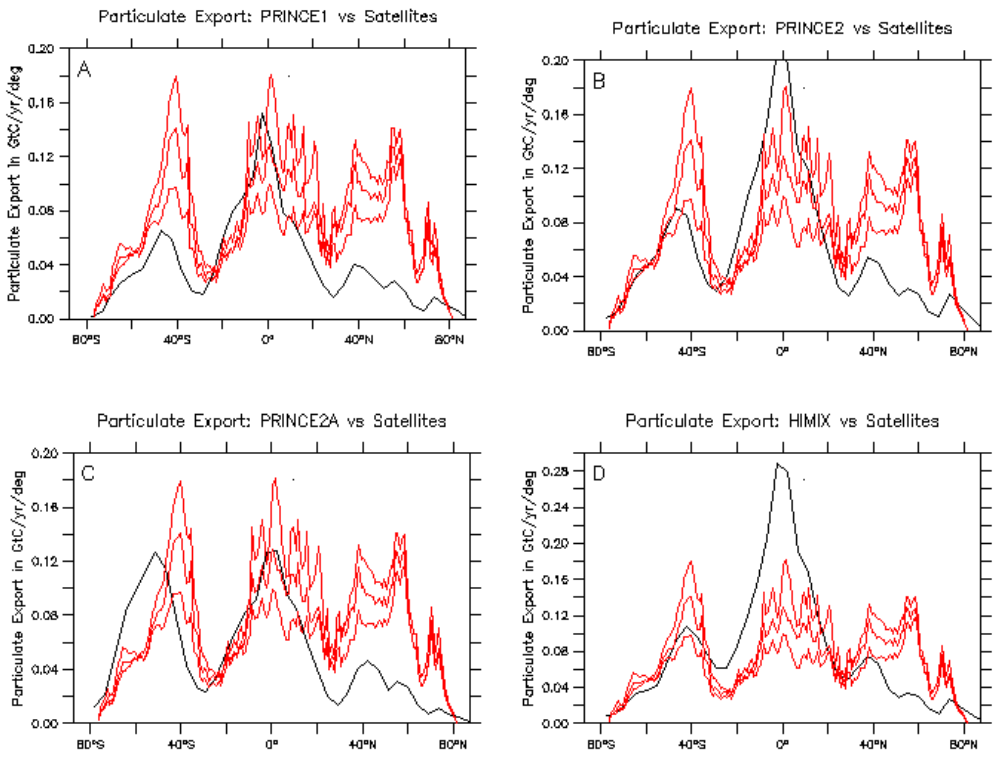


Figure 11: Comparison of observational estimates and models of particle export. Dashed lines are the observational estimates. Solid lines are the models. (A) PRINCE1. Note that this model is roughly in the right range in the tropics, but lies below the observational estimates outside the tropics. (B) PRINCE2. This model is significantly higher than observations in the tropics, close to the Marra curve in the Southern Ocean, and too low in the north. (C) PRINCE2A. Note the displacement of the Southern Ocean export in this simulation. (D) HIMIX. Vertical scale is expanded. Note that the tropical production is now severely overestimated.

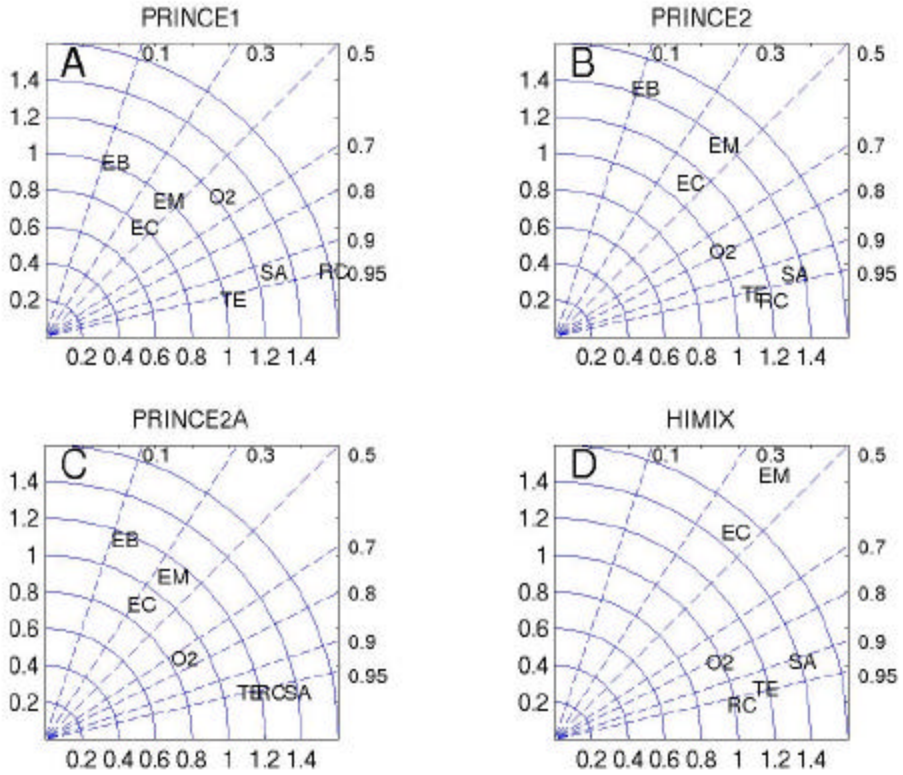


Figure 12: Taylor diagrams including particle export for the four models. TE is temperature, SA salinity, O2 oxygen, RC radiocarbon, EB export production using Behrenfeld and Falkowski (1997), EC export production using Carr (2002) and EM export production using Marra (in prep.). Export production is the monthly-varying zonal average. The results show graphically that although the newer models produce a tighter clustering of the hydrographic parameters, they do not produce a tighter clustering of the export production. (a) PRINCE1. (b) PRINCE2 (c) PRINCE2A (d) HIMIX (EB is off the chart for this model).